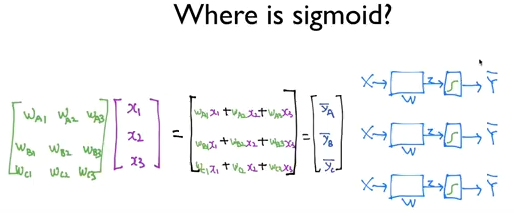
**6-2장. Softmax classifier의 cost 함수**

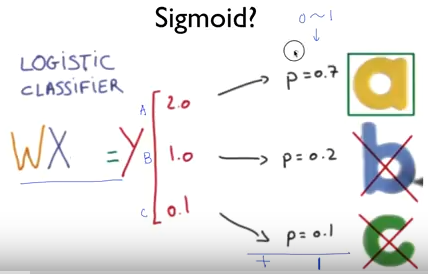


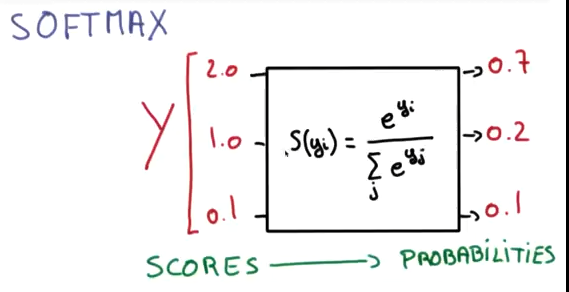
WX 매트릭스곱셈으로 계산된 값을 sigmoid함수에 통과하면 각각 0~1의 예측값 Y가 계산된다.

그런데, 예측값 Y 가 0~1 이면서 예측값의 합계가 1이 되도록 하는 함수가 있다면?

그렇다면 예측값 Y 를 A, B, C 가 나올 확률로 볼 수 있게 된다.

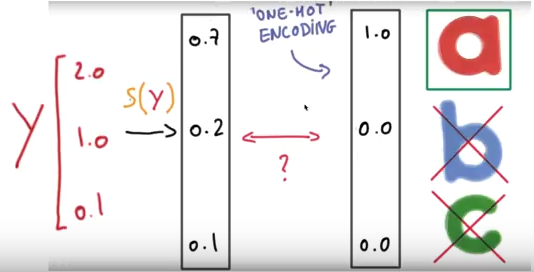
Softmax 함수가 바로 위의 두가지 조건(0~1 이고, 합계가 1)을 만족하는 함수이다.



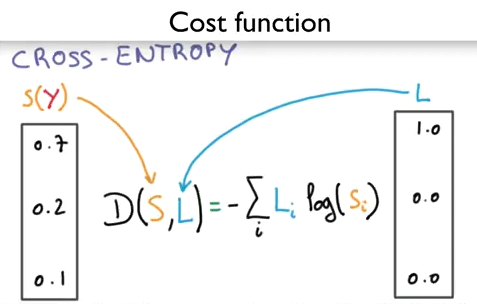


(\*) softmax 함수는 0~1 사이의 값이고, 전체의 합이 1이 되는 값을 계산하는 함수이다.

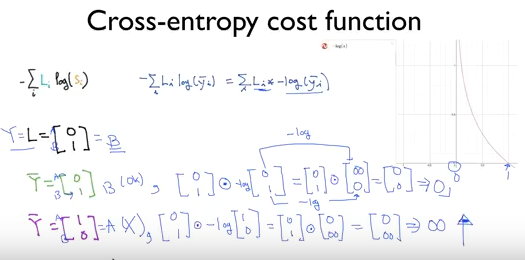
또한, 최종적으로 ONE-HOT ENCODING 이라는 기법을 사용하면 가장 큰 값은 1, 그외는 0을 리턴하도록 할 수도 있다.



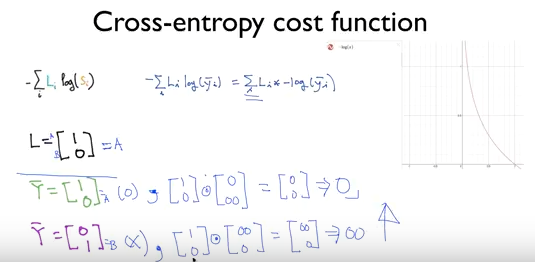
(\*) 실제값 L 과 예측값 S(y)의 차이인 Cost 함수는 Cross-entropy 함수를 사용하여 계산한다

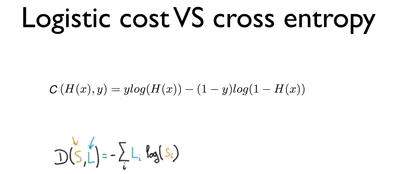


(\*) 실제값(L)이 B 일때, B로 예측했을때의 cost값은 0, A로 예측했을때의 cost값은 무한대가 된다.



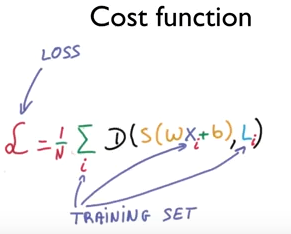
(\*) 실제값(L)이 A 일때, A로 예측했을때의 cost값은 0, B로 예측했을때의 cost값은 무한대가 된다.



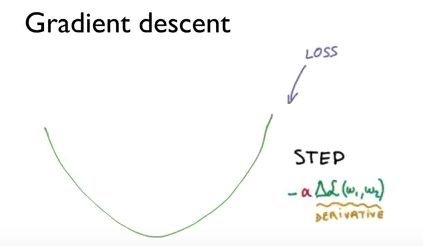


(\*) Logistic cost function 과 cross entropy cost function은 서로 다르지만

실제로 두 수식은 같은 것이다. 왜 같은 것인지 교수님은 숙제로 남겨주셨는데 잘 모르겠다?



(\*) 여러 개의 training set 이 있을때는 전체의 거리를 구한 후 합한 값의 평균을 계산하여 cost를계산한다.



(\*) 위의 cost function 그래프도 Gradient descent 알고리즘을 사용하여 cost값이 최소가 되는 W를 찾을수 있다.